A simple algorithm for optimization and model fitting: AGA (Asexual Genetic Algorithm)

J. Cantó, S. Curiel and E. Martínez-Gómez

Instituto de Astronomía, Universidad Nacional Autónoma de México, Apdo. Postal 70-264, Ciudad Universitaria, Coyoacán, 04510, Mexico

e-mail: scuriel@astroscu.unam.mx; affabeca@gmail.com

Abstract

Context. Mathematical optimization can be used as a computational tool to obtain the optimal solution to a given problem in a systematic and efficient way. For example, in twice-differentiable functions and problems with no constraints, the optimization consists of finding the points where the gradient of the objective function is zero and using the Hessian matrix to classify the type of each point. Sometimes, however it is impossible to compute these derivatives and other type of techniques must be employed such as the steepest descent/ascent method and more sophisticated methods such as those based on the evolutionary algorithms.

Aims. We present a simple algorithm based on the idea of genetic algorithms (GA) for optimization. We refer to this algorithm as AGA (Asexual Genetic Algorithm) and apply it to two kinds of problems: the maximization of a function where classical methods fail and model fitting in astronomy. For the latter case, we minimize the chi-square function to estimate the parameters in two examples: the orbits of exoplanets by taking a set of radial velocity data, and the spectral energy distribution (SED) observed towards a YSO (Young Stellar Object).

Methods. The algorithm AGA may also be called genetic, although it differs from standard genetic algorithms in two main aspects: a) the initial population is not encoded, and b) the new generations are constructed by asexual reproduction.

Results. Applying our algorithm in optimizing some complicated functions, we find the global maxima within a few iterations. For model fitting to the orbits of exoplanets and the SED of a YSO, we estimate the parameters and their associated errors.

Key words. Methods:numerical–Stars–55 Cancri–Planets and Satellites:general–ISM:individual(L1448)

1. Introduction

Mathematical optimization can be used as a computational tool in deriving the optimal solution for a given problem in a systematic and efficient way. The need to search for parameters that cause a function to be extremal occurs in many kinds of optimization. The optimization techniques fall in two groups: *deterministic* (Horst and Tuy, 1990) and *stochastic* (Guus, Boender and Romeijn ,1995). In the first group, we have the classical methods that are useful in finding the optimum solution or unconstrained maxima or minima of continuous and twice-differentiable functions. In this case, the optimization consists of identifying points where the gradient of the objective function is zero and using the Hessian matrix to classify the type of each point. For instance, if the Hessian matrix is positive definite, the point is a local minimum, if it is negative, the point is a local maximum, and if if indefinite, the point is some kind of saddle point. However, the classical methods have limited scope in practical applications since some involve objective functions that are not continuous and/or not differentiable. For these reasons, it is necessary to develop more advanced techniques that belong to the second group. Stochastic models rely on probabilistic approaches and have only weak theoretical guarantees of convergence to the global solution. Some of the most useful stochastic optimization techniques include: adaptive random search (Brooks, 1958), clustering methods (Törn, 1973), evolutionary computation that includes genetic algorithms, evolutionary strategies and evolutionary programming (Fogel, Owens and Walsh, 1966; Schwefel, 1995; Goldberg, 1989; McCall, 2005), simulated and quantum annealing (Kirkpatrick, Gellatt and Vecchi, 1983), and neural networks (Bounds, 1987).

We present a simple algorithm for optimization (finding the values of the variables that maximize a function) and model fitting (finding the values of the model parameters that fit a set of data most closely). The algorithm may be called genetic, although it differs from standard genetic algorithms (Holland, 1975) in the way that new generations are constructed. Standard genetic algorithms involve *sexual reproduction*, that is, the reproduction by the union of male and female reproductive individuals. Instead, our algorithm uses *asexual reproduction*, in which offspring are produced by a single parent (as in the fission of bacterial cells).

The paper is organized as follows. In Sect. 2, we present and describe the main characteristics of our algorithm called AGA (**A**sexual **G**enetic **A**lgorithm). In Sect. 3, we apply the algorithm to two kinds of problems: maximization of complicated mathematical functions and a model fitting procedure. In the latter group, we consider two examples taken from astronomy: a) the orbital fitting of exoplanets, and b) the model fitting of the Spectral Energy Distribution (SED) observed in a Young Stellar Object (YSO). In both cases, we minimize their corresponding chi-square function. In Sect. 4, we summarize and discuss the results for each case.

2. Description of the AGA (Asexual Genetic Algorithm)

We consider the problem of finding the absolute maxima of a real function of N variables i.e., identifying the values of the N variables (the coordinates of a point in the space of N dimensions) for which the function attains its maximum value. It is assumed that the absolute maximum is inside a bounded region V where the function is defined.

Our algorithm proceeds in the following way (see also Fig. 1):

1. Construct a random initial population. The initial population is a set of N_0 randomly generated points (in the context of evolutionary algorithms, they are also called individuals) within the region V.
2. Calculate the fitness of each individual in the population. The fitness is calculated by evaluating the function at each point.
3. Select a subset of individuals with the highest fitness. Rank the points according to the value of the function, and choose a subset of N_1 points with the highest values of the function.
4. Construct a new population using the individuals in the subset. Generate N_2 random points within a previously selected vicinity E around each of the selected points.
5. Replace the source population with the new population. The new population is the set of $(N_1 \times N_2) + N_1$ points that results from step 4 plus a clone of each parent. We may choose (as we did) N_1 and N_2 such that $(N_1 \times N_2) + N_1 = N_0$, keeping the size of the population N_0 unchanged for each generation. In this way, one can devise an iterative procedure.
6. If the stopping criteria (accuracy, maximum number of generations, etc.) have not been met, return to step 2.

It is clear that this presented algorithm in many aspects resembles standard genetic algorithms. The key difference is the way the new population is constructed: in this version, we propose an *asexual reproduction with mutation* in the sense that the offspring of a parent is a point randomly selected within a narrow neighborhood of the (single) parent point. Hence, the name for our algorithm is AGA, meaning **A**sexual **G**enetic **A**lgorithm (Fig. 1). We note that in the new generation we always include a clone of the parent, that is, we always keep the original seed points to be used for the next generation whenever the parents are more well suited than their offspring.

For the vicinity E of each selected point, we used rectangular (hyper)boxes of decreasing size. The box around each point was centered on the point and had a side length $2\Delta_{ij}$ along direction i in the generation j. In particular, we take,

$$\Delta_{ij} = \Delta_{i0}p^j, \tag{1}$$

where $2\Delta_{i0}$ is the initial length of the box along direction i, and p is a fixed numerical value less than unity (which can be called the "convergency factor"). In this way, the length of each box side decreases by a factor p in each generation. For instance, if we want the side length of the box to decrease by a factor 2 after 10 generations then,

$$p = \left(\frac{1}{2}\right)^{\frac{1}{10}} = 0.9330 \tag{2}$$

Decreasing the size of the vicinity E each generation is intended to achieve the highest possible accuracy for the position of the point at which the function attains its absolute maxima.

The speed with which the AGA finds a solution depends, of course, on the factor p; the lower the value of p, the faster the solution found. However, if p is too low, the sampling area around the points may decrease so fast that the AGA has no time to migrate to the true solution. On the other hand, if p is too high, many of the offspring will never reach the solution within the convergency criterion. As a consequence of this, the error in the solution will be high and in some cases, the solution will not be reached. The adequate value for p depends on the problem itself. The optimization of p can be achieved by trial and error. However, we have found that a value between 0.4 and 0.6 is adequate for all the tested problems.

An alternative way of choosing the box size consists of employing the standard deviation of the points contained in the subset N_1 along each dimension. In such a case, the length of the sides of the box naturally decreases as the algorithm converges. This method works quite efficiently for problems with a few dimensions. Interestingly, the length of the box size in this case decreases following a power law such as that in Eq. 1, with a moderately high values of p (between 0.5-0.8) for the first few generations, abruptly changing to a much lower values (0.2-0.3) for the rest of the generations.

In the case of problems with the large number of parameters to be estimated, we added an iterative method to the scheme presented above. This iterative method consisted of performing a series of runs (each one following the scheme shown in Fig. 1) in such a way that the resultant parameters of a run are taken as the initial "guess" for the next run. This procedure may be equivalent to performing a single run allowing for additional generations, but this is not the case. The key difference is that for each run in the iterative procedure, the sizes of the sampling boxes are reset to their initial values, i. e., each run starts searching for solution using boxes of the same size as those used in the first run but centered on improved initial values. We consider to have reached the optimal solution when the values of the parameters do not change considerably (within a tolerance limit) after several iterations, and the χ^2 is found to have reached a limiting value (see sect. 3.2.1 and Fig. 7 for an example).

We find that this iterative strategy guarantees the convergence to the optimal solution, since it avoids the potential danger of using values of p that do not allow the AGA to drift (migrate) to the "true" solution within a single run. Problems that involve the finding of a large number of parameters are potentially subject to this risk, since each parameter may require a different value of p. Furthermore, problems with a large number of parameters become particularly difficult when the values of the different parameters differ by several orders of magnitude, as in the fitting of the orbits of exoplanets and the fitting of the SED of YSOs (see Examples 1 and 2 in Sect. 3). In this case, the iterative method has proven to be particularly useful; the solution usually improves considerably after several iterations.

In the following section, we describe some applications of AGA. We have divided the applications into two groups depending on the type of problem to solve: the maximization of complicated functions and model fitting in astronomy.

3. Applications of the AGA

We separate the optimization problems in two groups. In the first group, we consider functions of two variables where many classical optimization methods present formidable difficulties in finding the global maximum. In the second group, we show two typical examples of model fitting in astronomy that can be treated as minimization procedures.
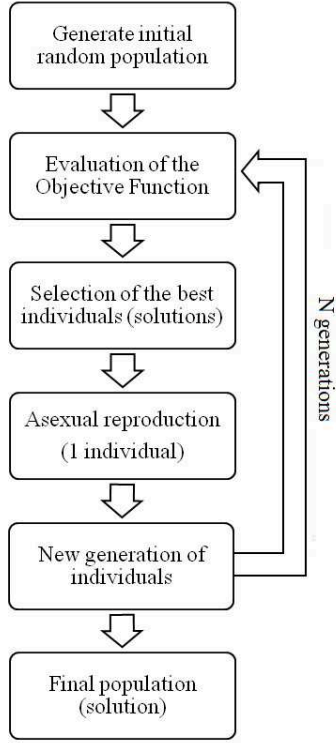
Figure 1. Basic diagram for the implementation of the Asexual Genetic Algorithm (AGA). First we generate a random initial population. Then, we evaluate the fitness of each individual in this population and select those which have the highest fitness. A new generation is constructed by an asexual reproduction (see text) which replaces the older one. If the stopping criteria are met, we stop. If not, we use these individuals as an initial population and start again.

3.1. Optimization of functions of two variables

There are many examples of functions that are not easy to optimize with classical techniques such as the simplex method, the gradient (or Newton-Raphson) method, the steepest ascent method (Everitt, 1987), among others. In such cases, the existence of many maxima (or minima) and the sharpness of the peaks can represent a serious problem. Because of this, the standard Genetic Algorithms or GA are successfully applied in searching for the optimal solution. We consider some typical examples treated by this technique, which are shown below.

3.1.1. Example 1

We consider the following function (Charbonneau, 1995):

$$f(x,y) = [16x(1-x)y(1-y)\sin(n\pi x)\sin(n\pi y)]^2 \quad (3)$$

where the variables x, $y \in [0,1]$ and $n \in \mathbb{N} - \{0\}$. Identifying the global maximum of this function for large n is a difficult task because there are many local maxima that differ little in value but are separated by deep "valleys" in the two-dimensional landscape. Techniques such as the steepest ascent/descent and the conjugate gradient method are local methods that work well if $f(x, y)$ is a smooth function that can be differentiated at least once and a single maximum exists in the domain under consideration.

Table 1. Values of the initial parameters used in AGA for the four examples presented in this work.

AGA's parameter / Example	p	N_1	N_2	N_0
Bi-dimensional functions (2)	0.6	10 20	9 19	100 400
Extrasolar planets YSO	0.45	30	29	900

In Fig. 2, we show the solution to this problem ($n = 9$) by applying AGA after fixing the initial population size, number of parents, number of descendants, and convergency factor (Table 1). Our graphs have the same format as those presented by Charbonneau (1995) to facilitate their comparison.

In the first panel, we start with a population of $N_0 = 100$ individuals, i. e., a set formed by 100 random points representing the candidate solutions to the global maximum distributed more or less uniformly in parameter space. After 5 generations (second panel), a clustering at the second, third and fourth maxima is clearly apparent. After 10 generations, the solutions already cluster around the main maxima. At the 25th generation, we have reached the maximum in $(x^*, y^*) = (0.5, 0.5)$ with $f(x^*, y^*) = 1$ with an accuracy of $\sim 10^{-7}$. We note that at the 25th generation all the 100 individuals have reached the maximum with at least this accuracy.

In the last panel, we show the evolution in the fittest *"phenotype"* with the number of generations as plotted for two sizes of the population, $N_0 = 100$ and $N_0 = 400$; in other words, we measure the deviation in the function value for the maxima points identified in each generation from the "true" maximum, that is, $f(0.5^*, 0.5^*) = 1$. It is evident from Fig. 2 that a larger size of the population causes the maximum to be reached in a lower number of generations.

In Fig. 3, we show the evolution in the global solution using AGA and the results obtained by Charbonneau (1995) using a traditional GA. In the first panel for both works, an initial population of 100 individuals is randomly selected. After 10 generations, all the individuals in our algorithm have clustered around the maxima; in contrast, the GA has started to form some clusters. In the third panel, after 20 generations, AGA has found the global maxima, while the GA continues to search among clusters for a global solution. In the fourth panel, the convergence to the global solution is shown for both works, our AGA having reached the solution in 25 generations with higher accuracy than the 90 generations employed by the GA. In the fifth panel, we finally can compare the evolution in the fittest phenotype for each generation, calculated to be $1 - f(x_{max}, y_{max})$. For AGA, these differences are much smaller than those obtained with the GA method, which indicates that our algorithm is more accurate than GA. However, we note that the solution presented in Charbonneau's work was limited by the number of digits used in the encoding scheme. The GA algorithm can also reach high levels of precision by changing the encoding, but at the cost of slower convergence, which would make the GA algorithm become even more computationally expensive than the AGA. The lower number of generations required by AGA to reach a global solution demonstrates that AGA is more efficient than the GA method.
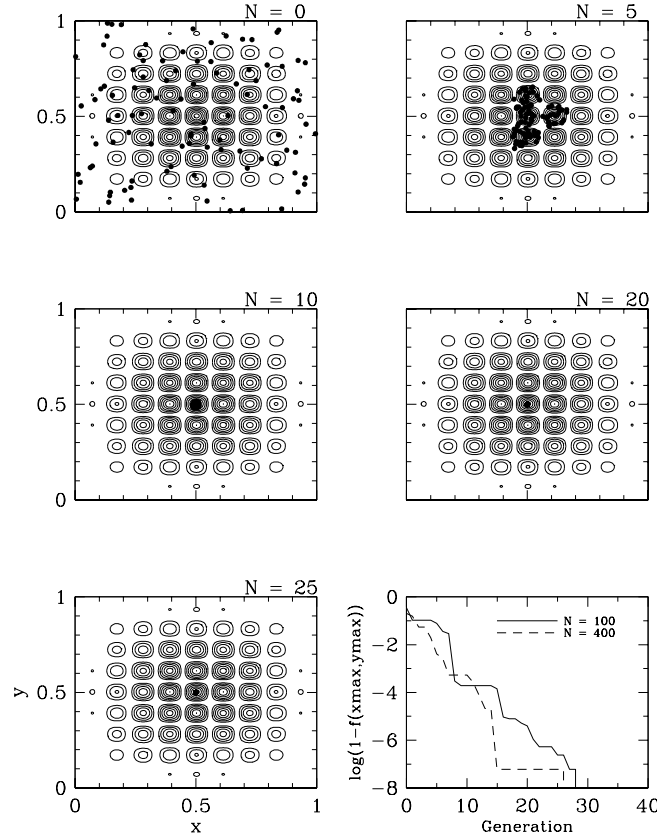
Figure 2. A solution to the model optimization problem based on the idea of the Asexual Genetic Algorithm. The first five panels show the elevation contours of constant f (Eq. 3 with $n = 9$) and the population distribution of candidate solutions (each one contains 100 points), starting with the initial random population (in the standard GA it is defined as the *"genotype"*) in (a) and proceeding on through the 25th generation on (e). In the sixth panel, we show the evolution of the fittest *"phenotype"* assuming two sizes of the population, $N_0 = 100$ and $N_0 = 400$.

3.1.2. Example 2

The following example is proposed as a function test in the PIKAIA's user guide (Charbonneau and Knapp, 1995). The problem consists of locating the global maximum of the function:

$$f(r) = f(x, y) = cos^a(b\pi r) \, exp\left(-\frac{r^2}{2\sigma^2}\right) \qquad (4)$$

where a and b are known constants, r is the radial distance given by the expression $r = \sqrt{(x - x_0)^2 + (y - y_0)^2}$, σ is the width of the Gaussian and the position of the maximum peak is (x_0, y_0). Observed from above, this function appears like concentric rings of similar widths and amplitudes (see Fig.4). The difficulty in finding the maximum peak of this function is that the area of the concentric rings is much larger than the area at the center containing the maximum peak (see Fig.4). This property ensures that most of the algorithms searching for the maximum peak fail because they usually become fixed inside in one of the rings and identify a local maximum instead. We experimented with different values of a, b, and σ^2 ($a = 2, 4$; $b = 3, 9$ and $\sigma^2 = 1, 2, 4$), as well as different centers (x_0, y_0). In all cases, using AGA with the parameters shown in Table 1, the "true" maximum peak was found in just a few steps, typically in less than 100 generations, and with an accuracy of 10^{-6}. The results

Table 2. Values of the parameters a, b, and σ^2 assumed for the function $f(r) = cos^a(b\pi r) \, exp\left(-\frac{r^2}{2\sigma^2}\right)$ and the number of generations needed to reach an error tolerance of 10^{-6}.

a	b	σ^2	Ngen
2	3	1	65
2	3	2	69
2	3	4	66
2	9	1	94
2	9	2	94
2	9	4	91
4	3	1	69
4	3	2	66
4	3	4	58
4	9	1	94
4	9	2	101
4	9	4	98

are summarized in Table 2. Changing the position of the center did not substantially change the number of generations needed to reach the selected tolerance level.
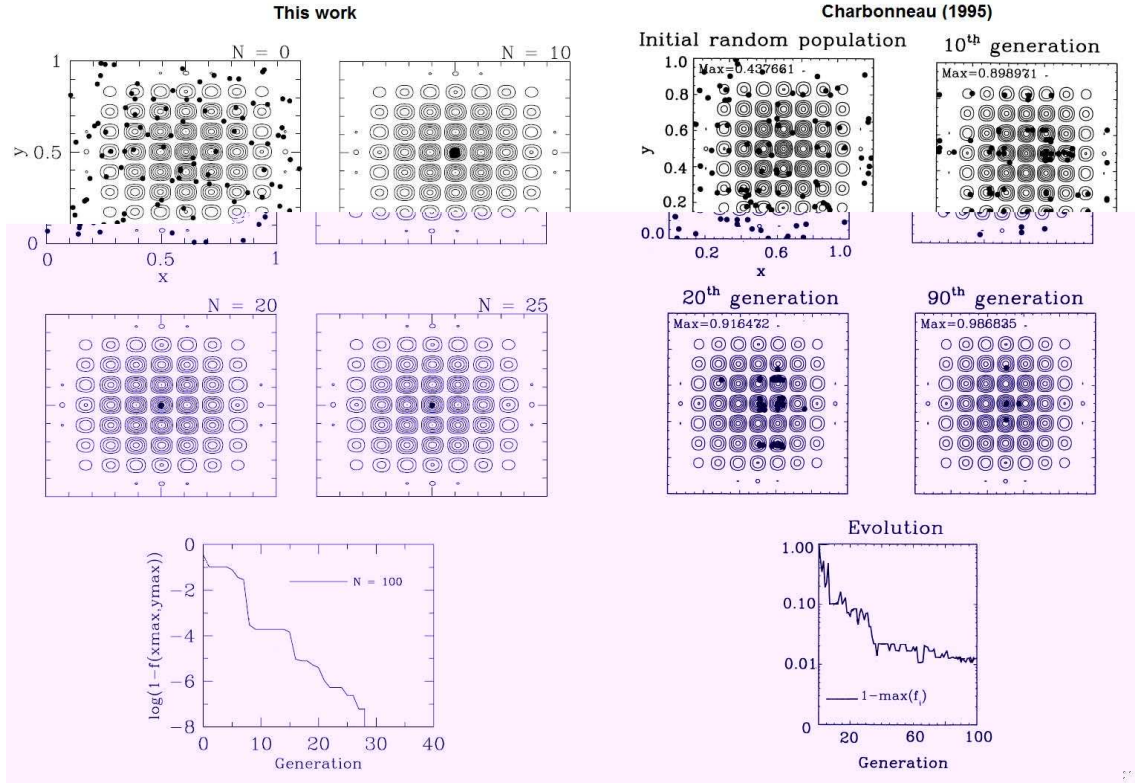
Figure 3. Comparison of our results with those obtained by Charbonneau (1995). We have maintained the same format and removed the curve of the evolution of the median-fitness individual in the fifth panel of Charbonneau's work to facilitate direct comparison. In the first panel we start with an initial population of 100 individuals and both algorithms start to search the global maxima. While the GA employs 90 generations to reach the solution, our AGA just requires 25 (fourth panels). The fifth panels show the evolution of the fittest phenotype with the number of generations for AGA (left) and GA (right). Note that indeed, AGA reaches an accuracy of $\sim 10^{-8}$ in less that 30 generations while GA attains an accuracy of $\sim 10^{-2}$ in 100 generations.

3.2. The parameter estimation problem in astronomical models

In the physical sciences, curve or model fitting is essentially an optimization problem. Giving a discrete set of N data points (x_i, y_i) with associated measurement errors σ_i, one seeks the best possible model (in other words, the closest fit) for these data using a specific form of the fitting function, $y(x)$. This function has, in general, several adjustable parameters, whose values are obtained by minimizing a "merit function", which measures the agreement between the data and the model function $y(x)$.

We suppose that each data point y_i has a measurement error that is independently random and distributed as a normal distribution about the "true" model with standard deviation σ_i. The maximum likelihood estimate of the model parameters $(c_1, ..., c_k)$ is then obtained by minimizing the function,

$$\chi^2 \equiv \sum_{i=1}^{N} \left(\frac{y_i - y(x_i; c_1, .., c_k)}{\sigma_i} \right)^2 \qquad (5)$$

3.2.1. Evaluation of the error estimation

The experimental or observational data are subject to measurement error, thus it is desirable to estimate the parameters in the chosen model and their errors. In the straight-line-data fitting and the general linear least squares, we can compute the standard deviations or variances of individual parameters through simple analytic formulae (Press et al., 1996). However, when we

attempt to minimize a function such as Eq. 5, we have no expression for calculating the error in each parameter. A good approach to solve this problem consists of building "synthetic data sets". The procedure is to draw random numbers from appropriate distributions so as to mimic our clearest understanding of the underlying process and measurement errors in our apparatus. With these random selections, we compile data sets with exactly the same numbers of measured points and precisely the same values of all control or independent variables, as our true data set. In other words, when the experiment or observation cannot be repeated, we simulate the results we have obtained.

We compiled the synthetic data with the process illustrated in Fig. **5** and from the following expression:

$$y_i' = y_i + \sigma_i (2\xi - 1) \qquad (6)$$

where y_i' represents the i-th data point of the new set, y_i is the i-th data point of the original set, σ_i is the error associated with the i-th data point, and ξ is a random number within the interval $[0, 1]$. Using Eq. 6, we generated synthetic data sets with the original errors because they are related to the measurements.

For each of these data sets, we found a corresponding set of parameters (see Fig. 5). Finally, we calculated the average values of each parameter and the corresponding standard deviation, as the estimates of the parameters and their associated errors, respectively. We applied this procedure to the following examples.
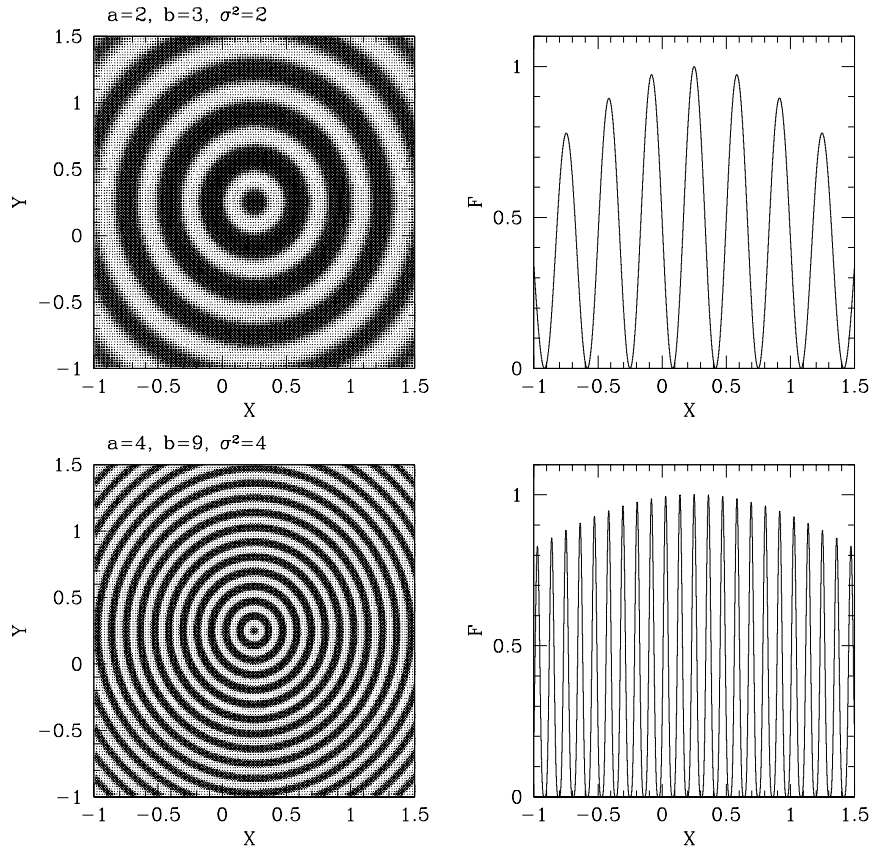
Figure 4. Profile of the bi-dimensional positive cosine function. In the upper panels, we take $a = 2$, $b = 3$ and $\sigma^2 = 2$; in the lower panels, we take $a = 4$, $b = 9$ and $\sigma^2 = 4$. In both cases, the position of the maximum peak is $(x, y) = (0.25, 0.25)$, where $f(0.25, 0.25) = 1$.
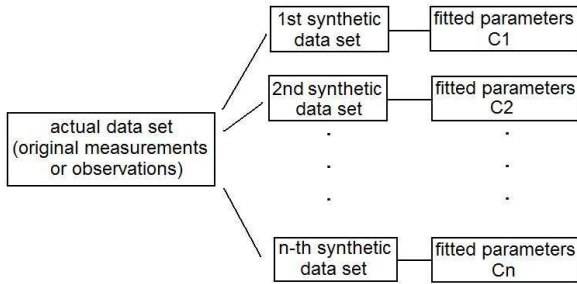

Figure 5. From the original set of measurements/observations, several synthetic data sets are constructed. This is achieved by adding to the dependant variable a random number whose absolute value is within the estimated errors of the original data. For each of these synthetic data sets, a set of fitted parameters is obtained. The average value of each parameter and the corresponding standard deviation are taken as estimates of the parameters and their associated errors.

3.2.2. Fitting the orbits of Extrasolar Giant Planets

We use the merit function given in Eq. 5 and the algorithm described in Sect. 2 to solve an interesting and challenging task in astronomy: curve or model fitting of data for the orbits of the extrasolar planets.

The first extrasolar planet was discovered in 1995 by Mayor and Queloz (1995) and, according to The Extrasolar Planets Encyclopaedia[1], until February 2009 there had been 342 candidates detected. Most of them (316) were revealed by radial velocity or astrometry of 269 host stars, that is, by Keplerian Doppler shifts in their host stars. Doppler detectability favors high masses and small orbits depending mainly on the present Doppler errors achievable with available instruments. In addition, the precision of the Doppler technique is probably about $3 \, \mathrm{m \, s^{-1}}$, owing to the intrinsic stability limit of stellar photospheres. This technique is sensitive to companions that induce reflex stellar velocities, $K > 10 \, \mathrm{m \, s^{-1}}$, and exhibit orbital periods ranging from a few days to several years, the maximum detectable orbital period being set by the time baseline of the Doppler observations. The remaining exoplanets were detected by other techniques: microlensing (8), imaging (11), and timing (7). For this example we only refer to the planets detected by radial velocity.

We now consider a system consisting of a central star of mass M_*, surrounded by N_p planets in bounded orbits. Assuming that the orbits are unperturbed Kepler orbits, the line-of-sight velocity of the star relative to the observer (Lee and Peale, 2003) is,

$$v_* = v_0 + \sum_{j=1}^{N_p} K_j \left[cos\left(f_j + \omega_j \right) + e_j \, cos \, \omega_j \right] \qquad (7)$$

[1] http://exoplanet.eu/catalog

where e_j is the eccentricity, ω_j is the argument of the pericenter, f_j is the true anomaly, K_j is the velocity amplitude of planet j, and v_0 is the line-of-sight velocity of the center-of-mass relative to the observer. The true anomaly is related to time by means of Kepler's equation,

$$E - e_j \sin(E) = \frac{2\pi}{T_j}(t - t_{0j}) \qquad (8)$$

and

$$\tan\left(\frac{f_j}{2}\right) = \sqrt{\frac{1 + e_j}{1 - e_j}} \tan\left(\frac{E}{2}\right) \qquad (9)$$

where T_j and t_{0j} are the period and time, respectively, of pericenter passage for planet j. Thus, for each planet, there are *5 free parameters*: e_j, T_j, t_{0j}, K_j and ω_j. Additionally, there is the systemic velocity v_0. In total, we have $(5 \times N_p) + 1$ free parameters, which are simultaneously fitted using AGA.

When these basic parameters for each planet are known through model fitting, we can estimate the semi-major axes of the orbits, a_j, and the masses of each planet, m_j. To do this, we have to make other simplifying assumptions. For instance, in the simplest model of totally independent planets,

$$K_j = \left(\frac{2\pi G}{T_j}\right)^{\frac{1}{3}} \frac{m_j \sin(i_j)}{(M_* + m_j)^{\frac{2}{3}}} \frac{1}{\sqrt{1 - e_j^2}} \qquad (10)$$

from which we can determine either m_j by assuming $\sin i_j = 1$ or $m_j \sin i_j$ by neglecting m_j in front of the mass of the star, M_*. We can then estimate a_j using Kepler's third law,

$$T_j^2 = \frac{4\pi^2 a_j^3}{G(M_* + m_j)} \qquad (11)$$

where G denotes the gravitational constant.

We applied the AGA presented in Sect. 2 to the measured radial velocities of the main-sequence star 55 Cancri, published by Fischer et al. (2008).

The data set contains 250 measurements completed at the Lick Observatory from 1989 to 2007, and 70 measurements made at the Keck Observatory from 2002 to 2007 of the velocity of 55 Cancri. Data from Lick has measurement errors of $\sim 10\,\mathrm{m\,s^{-1}}$ (1989-1994) and 3-5 $\mathrm{m\,s^{-1}}$ (1995-2007). Data from Keck has measurement errors of $3\,\mathrm{m\,s^{-1}}$ for data acquired prior to 2004 August and 1.0-1.5 $\mathrm{m\,s^{-1}}$ thereafter. The measured data have a maximum amplitude of $\sim 150\,\mathrm{m\,s^{-1}}$ and are of excellent signal-to-noise ratio. The orbital parameters were established by the detection of four planets in all Doppler measurements. To test a possible stellar systemic velocity remnant in the data, we added an additional parameter to the systemic velocity of the star obtaining a value of 17.28 $\mathrm{m\,s^{-1}}$. We fitted the orbits of 1, 2, 3, and 4 planets to this data (see Fig. 6). This exercise shows that the fitting solution improves when considering the orbit of more planets. The rms of the residuals and the χ_{red}, $\left(\chi_{red}^2 = \frac{\chi^2}{N_{points} - N_{parameters} - 1}\right)$ values, improve from values of 33.5 and 9.72, respectively, when fitting the orbit of one planet, to values of 7.99 and 2.03, respectively, when fitting the orbit of four planets (see Table 3). In general, the values of the rms, $\sqrt{\chi_{red}^2}$ fits and the derived parameters compare well with those obtained by Fischer et. al. (2008). In Tables 3 and 4, we summarize our results.

Table 3. Our values of the rms of the residuals and the $\sqrt{\chi_{red}^2}$ fits for the orbital fitting problem using AGA. For comparison, in the last two columns we show their corresponding values obtained by Fischer et al. (2008).

Planet [1]	This work rms (m s⁻¹)	This work $\sqrt{\chi_{red}^2}$	Fischer et al. rms (m s⁻¹)	Fischer et al. $\sqrt{\chi_{red}^2}$
b	33.5	9.72	39	10
b, c	10.69	3.38	11.28	3.42
b, c, d	9.69	2.49	8.62	2.50
b, c, d, e	7.99	2.03	7.87	2.12

[1] The single-planet fitting was obtained in Marcy et al. (2002).

With the exceptions of the third and fourth planet fittings, our rms values are lower than those obtained by Fischer et al. (2008). Our $\sqrt{\chi_{red}^2}$ values are however lower in all cases.

From the parameters shown in Table 4 we were able to derive the mass of the planet, $M_P \sin i$ (M_J), and the major semiaxis a (AU). For the four-planet fitting, we found (Table 5) that our values compare well with those reported by Fischer et al. (2008) in their five-planet model. Except for the first planet, all the values for the major semiaxis have a standard error lower than those obtained in the Fischer's model. In the case of the mass of each planet, our values have smaller standard errors.

We included the iterative scheme discussed in sect. 2 in fitting the orbit of the planets. In Fig. 7, we show the value of the χ_{red}^2 for the four-planet fit calculated after each run. In this case, we started with $\chi_{red}^2 = 4.16$ and after the 10th iterative run, the χ_{red}^2 had diminished to 4.15, which means that we had not found the optimal solution. During the first iterative runs, we note that the value of χ_{red}^2 decreased rapidly. After about 100 iterative runs, the χ_{red}^2 had not changed significally and converged to a fixed value (there are only slight variations in the last decimals). At this point, we can be assured that the value of the parameters have been found. Finally, we estimated the errors in fitting the orbits of the four planets as described in sect. 3.2.1. We present the estimated errors in Tables 4 and 5.

3.2.3. Fitting the spectral energy distribution (SED) for a YSO

A spectral energy distribution (SED) is a plot of flux, brightness or flux density versus frequency/wavelength of light. It is widely used to characterize astronomical sources i. e., to identify the type of source (star, galaxy, circumstellar disk) that produces these fluxes or brightness. Modelling the observed SEDs can help us to infer the temperature and size, among the other physical parameters of the source. As examples in radio astronomy, a SED with a negative spectral index ~ -0.7 would indicate the presence of a synchrotron radiation source; in infrared astronomy, SEDs can be used to classify T-Tauri stars; in galactic astronomy, the analysis of the SEDs leads to the determination of the respective roles of the old and young stellar populations in dust-grain heating.

For the reasons explained above, it is interesting to find the adequate model for an observed SED. We consider the observations reported by Curiel et al. (2009) in the L1448 region. This cloud is part of the Perseus molecular cloud complex located at a distance of $\sim 300\,\mathrm{pc}$. We are interested in fitting the observed SED for a couple of reasons: it contains an extremely young and highly collimated bipolar outflow (Bachiller et al. 1990) and seems to be a site of very recent star formation based on some
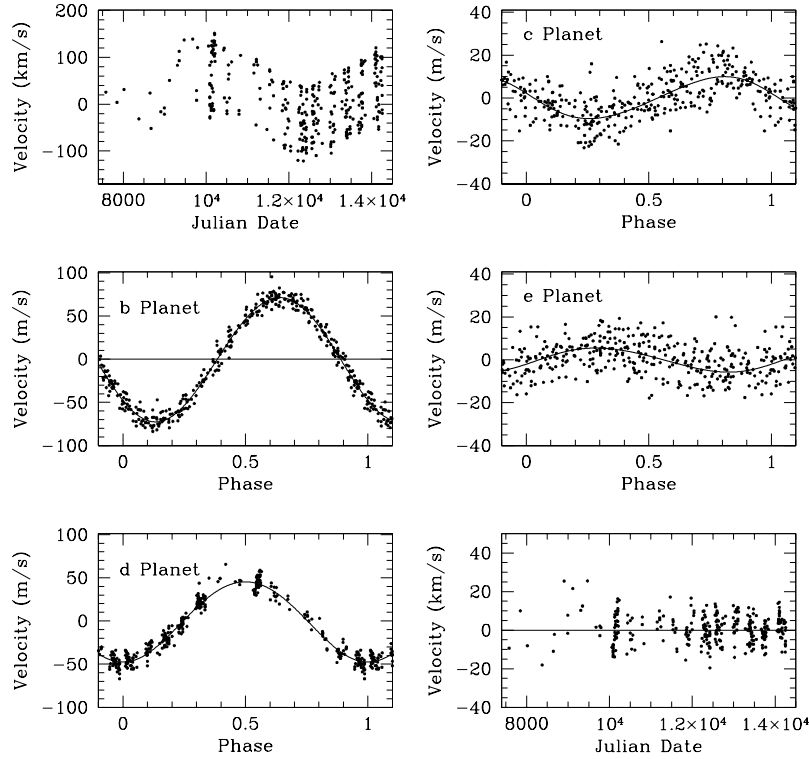
Figure 6. The velocities and fits for each of the four planets are shown separately for clarity by subtracting the effects of the other planets. That is, for planet labelled *b*, planets *c*, *d* and *e* have been removed from the data using the parameters found in the simultaneous four-planet fitting. For planet labelled *c* we have removed planets *b*, *d* and *e*. For planet labelled *d* we have removed planets *b*, *c* and *e*. For planet labelled *e* we have removed planets *b*, *c* and *d*. We have added the fitted systemic velocity for the star, which value is 17.2826 m s^{-1}. The last panel (bottom and right) shows the residuals between the data and the model. The first (up and left) panel shows the raw data.

Table 4. Our estimated values for the five parameters of the four exoplanets around 55 Cancri. The planets are listed in order of increasing orbital period, and the planet designations, b-e, correspond to the notation given by Marcy et al. (2002) and Fischer et al. (2008). The value for the $\sqrt{\chi_{red}^2}$ is also included.

Planet	T (days)	t_0 (JD)	e	ω (deg)	K_1 (m s^{-1})	$\sqrt{\chi_{red}^2}$
e	2.8170 ± 0.0932	10000.0046 ± 0.3250	0.07 ± 0.0016	250.2326 ± 0.4613	5.4311 ± 0.2241	
b	14.6515 ± 0.0002	10002.8917 ± 0.0549	0.0145 ± 0.0005	130.9176 ± 0.4591	71.7606 ± 0.3140	2.0314
c	44.3298 ± 0.0059	9989.9237 ± 0.4637	0.0853 ± 0.0014	78.2384 ± 0.4662	9.9820 ± 0.2075	
d	5218.3339 ± 0.5246	12500.7572 ± 0.3713	0.0250 ± 0.0006	180.2123 ± 0.2645	46.6872 ± 0.1431	

Table 5. Our derived values for the mass of the planet and the major semiaxis for the four-planet fitting. For comparison, we also show the values obtained by Fischer et al. (2008) for their five-planet model, where we have removed the fifth body.

Planet	This work $M_P \sin i$ (M$_J$)	This work a (AU)	Fischer et al. $M_P \sin i$ (M$_J$)	Fischer et al. a (AU)
e	0.0361 ± 0.0014	0.0383 ± 0.0008	0.034 ± 0.0036	$0.038 \pm 1.0 \times 10^{-6}$
b	0.8285 ± 0.0036	$0.1148 \pm 1.0 \times 10^{-6}$	0.824 ± 0.007	$0.115 \pm 1.1 \times 10^{-6}$
c	0.1661 ± 0.0035	$0.2402 \pm 2.1 \times 10^{-5}$	0.169 ± 0.008	$0.240 \pm 4.5 \times 10^{-5}$
d	3.8201 ± 0.0117	5.7705 ± 0.0004	3.835 ± 0.08	5.77 ± 0.11

observations (Anglada et al. 1989; Curiel et al. 1990; Barsony et al. 1998; Girart & Acord 2001; Reipurth et al. 2002).

The data are taken from Curiel et al. (2009), who carried out a fit of the SED by assuming that there are three main components contributing to the flux at different wavelengths. Thus, the fitting function was given by the contribution of an optically free-free component and two grey bodies:
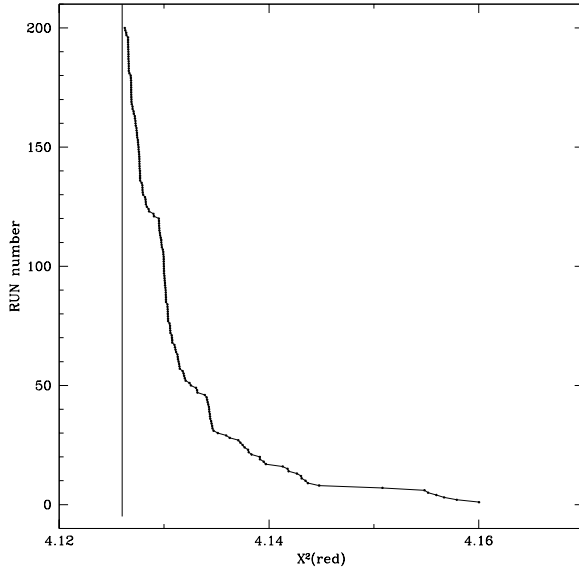
Figure 7. Convergence of the χ^2_{red} value with the number of runs (do not confuse with the number of generations done within AGA). The curve shows how χ^2_{red} converges to a "limiting" value in the sense that the variations occur in the last decimals (vertical line). As the number of runs increases the value of χ^2_{red} diminishes which means that we have not reached the optimal solution. Only after about 100 runs the χ^2_{red} value has converged and at this point the parameters do not significally change after subsecuent runs.

$$S_\nu = c_1\omega^{c_2} + \frac{c_3\left(1 - e^{-c_4\omega^{c_5}}\right)\omega^3}{e^{c_6\omega} - 1} + \frac{c_7\left(1 - e^{-c_8\omega^{c_9}}\right)\omega^3}{e^{c_{10}\omega} - 1}$$
(12)

where $\omega \equiv \frac{\nu}{\nu_0}$ is the frequency normalized to a reference frequency, ν_0.

The free parameters to be estimated, by the minimization of the chi-square (Eq. 5), are identified as $c_1, c_2, ..., c_{10}$. They are defined in terms of their corresponding physical parameters as: $c_1 \equiv F_C$ is the flux of optically thin emission with a spectral index given by $c_2 \equiv \alpha$, which corresponds to the free-free emission coming from a thermal jet; $c_3 = F_1 \equiv \frac{2h\Omega\nu_0^3}{c^2}$ is a reference flux; $c_4 \equiv \tau_1$ is the dust opacity evaluated at the reference frequency ν_0; $c_5 \equiv \beta_1$ is the dust emissivity index; and $c_6 \equiv \frac{h\nu_0}{kT_1}$ is related to the temperature. The last four parameters correspond to the emission originating in the first grey body (probably a molecular envelope). Finally, we propose that a second grey body exists that is a circumstellar disk surrounding the young protostar. Similarly, the parameters c_7, c_8, c_9, and c_{10} are related to the physical parameters of the second grey body.

In the previous expressions, ν is the frequency at which the source is observed, h is the Planck's constant, Ω is the solid angle subtended by the source, c is the speed of light, τ is the optical depth at the frequency ν_0 for each component, β is the dust emissivity index for each component, and k is the Boltzmann's constant. We assumed the characteristic frequency of the source to be $\nu_0 = \frac{c}{\lambda_{7mm}}$ (Curiel et. al., 2009). Table 6 summarizes the values of the fitted parameters in the model described by Eq. 12 of the SED for a YSO in the L1448 region. Table 6 also includes the estimated errors in the fitted parameters, as described in sect. 3.3.1. Figure 8 shows the observed data together with the best fit.

Table 6. Estimated values for the ten parameters in the model of the SED for a YSO in the L1448 region.

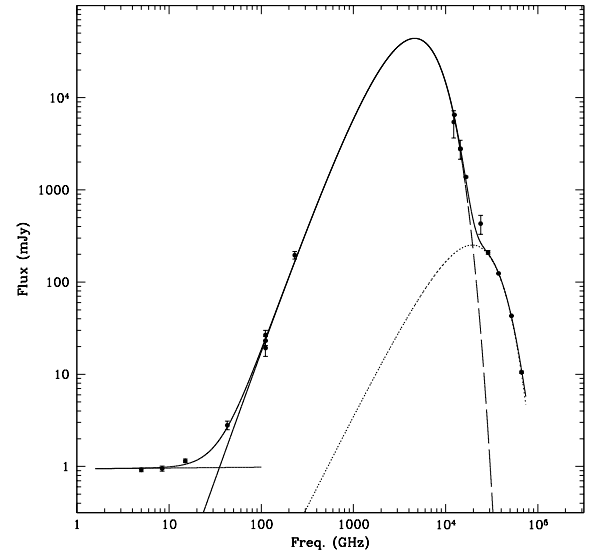| Parameter	Value	$\sqrt{\chi^2_{red}}$
F_C (mJy)	0.9295 ± 0.0330	
α	-0.0169 ± 0.0166	
F_1 (mJy)	1.7109 ± 0.1141	
τ_1	0.0728 ± 0.0062	
β_1	0.8999 ± 0.0575	
$T_{dust\,1}$ (K)	75.7522 ± 0.5273	1.9720
F_2 (mJy)	0.0068 ± 0.0008	
τ_2	34.9140 ± 40.7209	
β_2	1.5992 ± 0.3254	
$T_{dust\,2}$ (K)	330.8240 ± 1.2946	

Figure 8. The Spectral Energy Distribution (SED) for a YSO in the L1448 region. The dots represent the observations and the bars their associated measurement errors. The straight line corresponds to the continuum flux, the dashed line is the first grey body, the dotted line is the second grey body and the solid line is the sum of the three contributions.

From Fig. 8, we can conclude that the fitted model with the parameters shown in Table 6 is adequate for explaining the observations of the YSO in the L1448 region. The fitted curve lies within the measurement error bars (except one point located at high frequencies) and the SED consists of a continuum component and the emission from two grey bodies. Further astrophysical implications of these results can be found in Curiel et al. (2009).

4. Summary and Conclusions

We have presented a simple algorithm based on the idea of genetic algorithms to optimize functions. Our algorithm differs from standard genetic algorithms mainly in mainly two respects: 1) we do not encode the initial information (that is, the initial set of possible solutions to the optimization problem) into a string of binary numbers and 2) we propose an asexual reproduction as a means of obtaining new "individuals" (or candidate solutions) for each generation.

We have then applied the algorithm in solving two types of optimization problems: 1) finding the global maximum in functions of two variables, where the typical techniques fail, and 2) parameter estimation in astronomy by the minimization of the chi-square. For the latter case, we considered two examples: fitting the orbits of extrasolar planets associated with the star 55 Cancri and fitting the SED of a YSO for the L1448 region.

We found that our algorithm has several advantages:

– It is easy to implement in any computer because it does not require an encoding/decoding routine, and the new generations are constructed by the asexual reproduction of a selected subset (with the highest fitness) of the previous population.

– The algorithm does not require the evaluation of standard genetic operations such as crossover and mutation. This is replaced by a set of sampling rules, which simplifies the creation of new generations and speeds up the finding of the best solution.

– When the initial "guess" is far from the solution, AGA is capable to migrate searching for the "true" optimal solution. The final solution is usually achieved after a few hundred generations, in some cases, even faster.

– In some difficult cases, such as the fitting of the orbits of several exoplanets and the fitting of the SED of YSOs (with several components), AGA finds the solution in a few hundred generations. In these cases, an iterative scheme (several runs of AGA using the solutions of one run as the initial "guess" of the next run) can help to improve finding the "true" solution. This is particularly useful when the variables have different orders of magnitude, which causes the different variables to not converge to the final solution at the same time.

– As a consequence of the previous points, AGA becomes computationally less expensive than the standard version (GA). The convergence of AGA is reached in just a few generations.

From the example of the orbital fitting for exoplanets around 55 Cancri, we can conclude that our algorithm gives parameters that compare well with those obtained by Fischer et al. (2008). Marcy et al. (2002) also used the standard GA to obtain the parameters for a third planet, and they concluded that the GA fails to significantly improve the χ^2 value.

In the case of the SED for a YSO in the L1448 region, we find that the most adequate model contains the contribution of two grey bodies at high frequencies and free-free continuum emission, through a power law, at low frequencies.

Generally speaking the implementation of any kind of GA is advantageous in the sense that we do not have to compute any derivative of the selected fitness function allowing us to optimize functions with several local maxima points.

In this work, we have also implemented a method to estimate the error associated with each parameter based on the generation of "synthetic data sets". This method is easy to implement in any type of measured/observed data and constitutes a way of estimating the errors in the parameters when the minimization of the χ^2 is employed.

References

Anglada, G., Rodríguez, L.F., Torrelles, J.M., Estalella, R., Ho, P.T.P., Cantó, J., López, R., & Verdes-Montenegro, L. 1989, ApJ, 341, 208.
Bachiller, R., Martin-Pintado, J., Tafalla, M., Cernicharo, J. & Lazareff, B. 1990, A&A, 231(1), 174.
Barsony, M., Ward-Thompson, D., André, P., & O'Linger, J. 1998, ApJ, 509, 733.
Bounds, D. 1987, Nature 329, 215.
Brooks, S. H. 1958, Op. Res. 6, 244.
Charbonneau, P. 1995, Astrophys. J. Suppl. Series, 101, 309.
Charbonneau, P. & Knapp, B. 1995, A User's guide to PIKAIA 1.0, NCAR Technical Note 418+IA (Boulder: National Center for Atmospheric Research).
Curiel, S., Raymond, J. C., Moran, J. M., Rodríguez, L. F., & Cantó, J. 1990, ApJ, 365, L85.
Curiel, S., Girart, J. M., Fernández-López, M., Cantó, J. & Rodríguez, L. F. 2009, ApJ, submitted.
Everitt, B. S. 1987, in Introduction to optimization methods and their application to statistics, Chapman & Hall, Ltd, UK.
Fischer, D. A., Marcy, G. W., Butler, R. P., Vogt, S. S., Laughlin, G., Henry, G.W., Abouav, D., Peek, K. M. G., Wright, J. T., Johnson, J. A., McCarthy, C. & Isaacson, H. 2008, ApJ, 675, 790.
Fogel, L. J., Owens A. J. & Walsh, M. J. 1966, Artificial Intelligence through simulated evolution, Wiley & Sons, New York.
Girart, J. M., & Acord, J. M. P. 2001, ApJ, 552, L63.
Goldberg, D. E. 1989, in Genetic Algorithms in Search, Optimization and Machine Learning, Addison Wesley Longman Publishing Co, USA
Guus, C., Boender, E. & Romeijn, H. E. 1995, Stochastic methods. In Handbook of global optimization (eds. R. Horst and P. M. Pardalos), pp. 829-869, Kluwer Academic Publishers, Dordrecht, The Netherlands.
Holland, J. 1975, in Adaptation in Natural and Artificial Systems,
Horst, R. & Tuy, H. 1990, in Global optimization: Deterministic approaches, Springer-Verlag, Berlin.
Kirkpatrick, S., Gellatt, C. D. & Vecchi, M. P. 1983, Science, 220, 671.
Lee, M. H. & Peale, S. J. 2003, ApJ, 592, 1201.
Marcy, G. W., Butler R. P., Fischer, D. A., Laughlin, G., Vogt, S. S., Henry, G. W. & Pourbaix, P. 2002, ApJ, 581, 1375.
McCall, J. 2005, Journal of Computational and Applied Mathematics, 184, 205.
Mayor, M. & Queloz, D. 1995, Nature, 378, 355.
Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. 1996, in Numerical recipes in Fortran 77, Vol. 1, Cambridge University Press, UK.
Reipurth, B., Rodríguez, L. F., Anglada, G. & Bally, J. 2002, AJ, 124, 1045.
Schwefel, H. P. 1995, in Evolution and optimum seeking, Wiley & Sons, New York.
Törn, A. 1973, Global optimization as a combination of global and local search. Proceedings of computer simulation versus analytical solutions for business and economic models, pp. 191-206, Gothenburg, Sweden.